Post-Translational Modification

Introduction

Just as transcription had co-transcriptional modifications and post-transcription modifications, so too are there co-translational events and post-translational events. So far, we've led you to believe the growing amino acid strand is a straight line, a tail that just feeds off the ribosome and floats out in cytoplasm. It doesn't really do that.

Protein Folding

Proteins emerge from ribosomes and immediately start conforming into a three-dimensional shape. There are hydrophilic amino acids (they like cytoplasm) and there are hydrophobic amino acids (they don't like cytoplasm), so just from the structure alone there's going to be a propensity to hide the hydrophobic regions within the molecule or in the cell membrane.

The **primary structure** is simply the strand itself, the order in which the amino acids are assembled. Remember the code is 5' to 3', just as the amino acid sequence is N-term to C-term. Think of this as a line.

The **secondary structure** is the folding of that chain into one of two forms. Think of this as the line forming a plane, still in two dimensions. It can form into either an **α-helix** or a **β-pleated sheet**. There are other secondary structures (like collagen), but these are the ones that come up the most. The α-helix is like a cylinder of DNA and the β-sheet is essentially a two-dimensional plane.

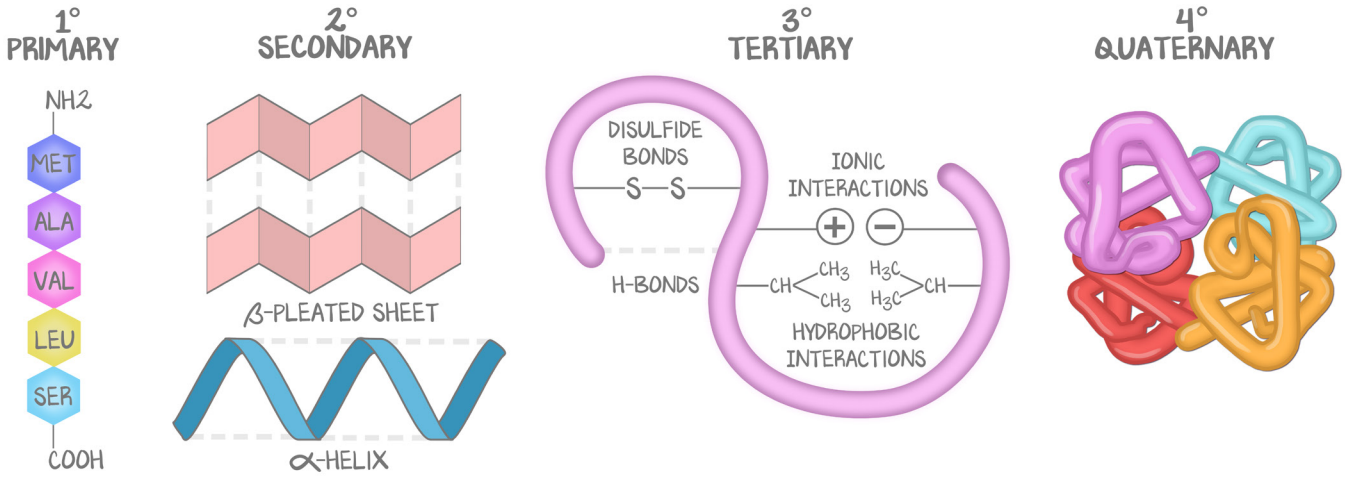

Figure 14.1: Protein Structure Level

The **tertiary structure** is the positioning of those structures. This is where it gets vague. There are so many interactions, so many different kinds of bonds in the protein structure, that it's impossible to predict. At this point, though, the structure is taking 3-dimensional form. When a protein is denatured, as by superheating, it's the tertiary structure that is compromised.

The **quaternary structure** is a combination of multiple subunits. Each subunit is a 3-dimensional protein and has its own tertiary structure. And the subunits come together. Like in hemoglobin, the GABA channel, or even the octamer histone.

The cell is constantly breaking down proteins and recycling the pieces. As new proteins get made, they're vulnerable until they're in their final form. Therefore, there are special classes of proteins called **chaperones** that protect the newly synthesized strand while the structure is obtained, and actually facilitate the structure itself. They prevent the formation of **aggregates** and prevent **misfolding**.

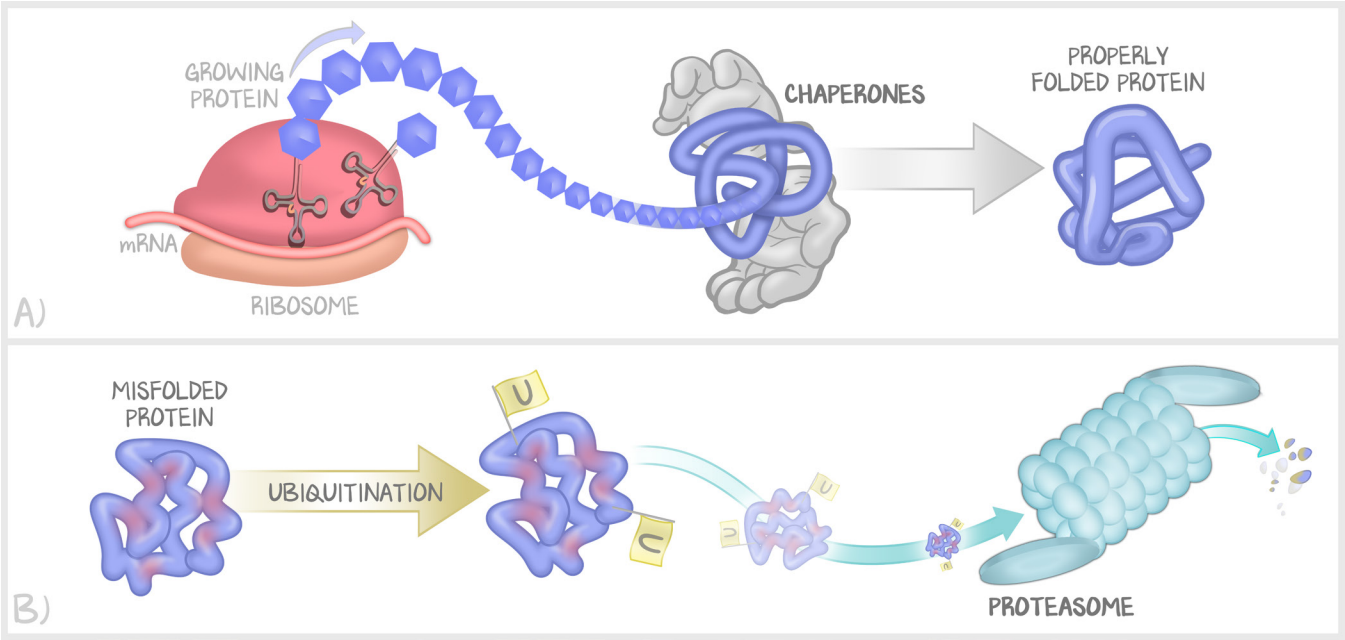

Figure 14.2: Ubiquitin and Proteasomes
Ubiquitin tags proteins for destruction by proteasomes. While new proteins are being made, chaperones protect the nascent amino acid chains from tagging with ubiquitin and facilitates appropriate folding. Improper folding or aggregation leads to ubiquitin tagging as a result of chaperone failure.

If a protein gets **misfolded**, aggregates, or for whatever reason doesn't get a chaperone, **ubiquitin** tags the protein for destruction. Where chaperones tell the cell, "this guy's all right, let him live," the ubiquitin (so called because it's ubiquitous, a function in all cells) tells the cell, "this guy's sick, get 'im." **Proteasomes** are enormous cytoplasmic complexes that effectively gobble and destroy any protein marked by ubiquitin.

Post-Translational Modification: Peptide Cleavage

The entire amino acid sequence isn't always used in the final function of a protein. For example, insulin is made in β-cells of pancreatic islets. Its primary amino acid sequence is **folded** such that the N-terminus and C-terminus come near to each other. They form **disulfide bonds**. That folded-on-itself-disulfide-bonded version of insulin is **proinsulin**. Proinsulin becomes insulin by **cleavage of the C-peptide**. The N and C termini are held together in a tertiary structure, and then the linking C-peptide is removed. The **functional insulin is made by cleaving away a large portion of the amino acids** (proinsulin → cleavage → insulin).

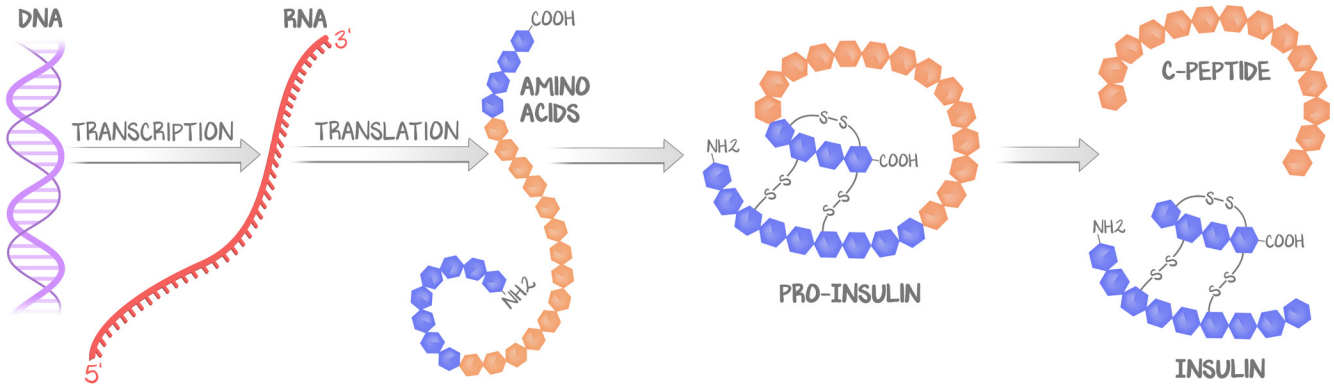

Figure 14.3: Insulin Peptide Cleavage
Insulin is transcribed into mRNA, then translated into a linear string of amino acids. It forms a tertiary structure by folding, and is stabilized by disulfide bonds (pro-insulin). The final product is released after peptide cleavage of the C-peptide into insulin.

Destination Distribution

Some proteins aren't destined for the cytoplasm. This could include any number of other proteins, such as secreted immunoglobulins, a cell-to-cell ligand, surface-bound membrane channels, those that go into lysozymes, etc. To get to anywhere other than the cytoplasm, these proteins must be in a thing that can merge with membranes. Membranes are lipophilic / hydrophobic. And once a protein is made, it's going to be quite the challenge to force a hydrophilic region through a lipophilic layer. Thus, it's better to start with the protein already in position for where it's supposed to go.

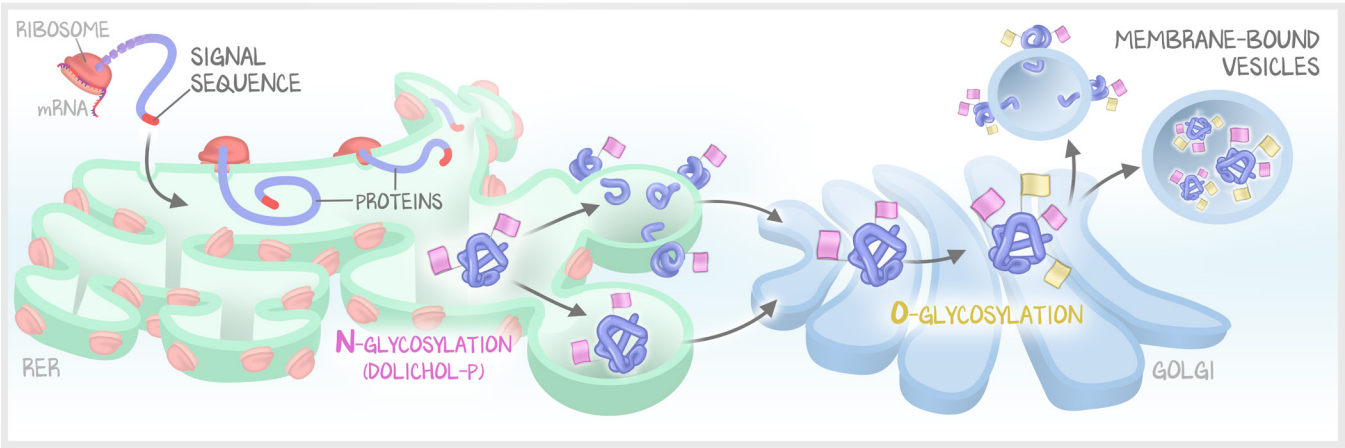

Figure 14.4: Signal Sequences

That means using the endoplasmic reticulum. As the ribosome is trucking down the mRNA, the **N-terminal hydrophobic signal sequence** marks the protein for the endoplasmic reticulum. So the entire complex—mRNA, rRNA, growing strand—heads towards the ER. As more and more ribosomes join, the smooth ER gets littered with ribosomes; it's now called an active, Rough Endoplasmic Reticulum (RER). The ribosomes are building protein from the mRNA, spilling the next amino acid into the RER. Some proteins will have hydrophilic ends on either side (receptors, channels), some will be unidirectional, and others will complete all of their processing contained by a lipid membrane (lysosomes, exocytosis).

Translational **N-glycosylation** occurs via **dolichol phosphate**. This marks the protein to be freely moved from the RER to the Golgi. In the Golgi, post-translational **O-glycosylation** can occur. Glycosylation goes to certain amino acids and has a specific purpose, but the level of detail needed isn't more than that. The point is that proteins that get tagged go into organelles, organelles then can become vesicles, vesicles can then combine with membranes and other membrane-bound structures.

It's the **hydrophobic N-terminal signal sequence** that starts the whole process. Once in the ER, **signal peptidase** removes the signal sequence. Translation continues as the ribosomes attached to the ER make it RER. **Glycosylation** (of any kind—N, O, etc.) occurs in the RER and/or Golgi. Proper folding is required to leave the RER to the Golgi. If going to a **lysosome**, there's **phosphorylation of mannose** by **phosphotransferase**.

The problem, though, is that proteasomes are out in cytoplasm. If a protein misfolds in the organelle, there isn't a good way to handle it. Some genetic disorders lose the ability to process these proteins at all, leaving a massive amount of protein accumulation stuck in an organelle that doesn't know what to do with it (i.e., α-1 antitrypsin deficiency with its PAS+ macrophages in the liver).

To Secrete or to Lysosome? That Is the Question

Lysosomal enzymes are glycosylated in a very specific way. If you want to secrete something, DON'T phosphorylate the N-linked oligosaccharides. When a protein arrives in the Golgi, it's marked for lysosome location by **n-acetyl-glucosamine-1 phosphotransferase** (gunner vocab word only) forming a mannose-6-phosphate. It's the **phosphorylation** that marks the protein as "lysosomal."

Summary

Glycosylation occurs through the ER and Golgi—O-glycosylation, N-glycosylation or both. Proteolysis can occur, breaking peptide bonds in order to restructure or reorient them to their final form. Phosphorylation, in particular for the lysosomal destination, happens in the Golgi.